


[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent &amp; Trademark Office

Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: **[((database or data <near> mining) and (sequence or gene or dna or nucleotides or biopolymer) <near> analysis and (subset or set or dataset) <paragraph> (reduce or eliminate) <near> (duplicate or redundant))]**  
 Found **236** of **126,861** searched.

Warning: Maximum result set of 200 exceeded. Consider refining.

Search within Results

129/865,090



&gt; Advanced Search

&gt; Search Help/Tips

Sort by: Title Publication Publication Date Score

Results 1 - 20 of 200

short listing

Prev  
Page

1 2 3 4 5 6 7 8 9 10

Next  
Page

# 1 Special issue on prototypes of deductive database systems: The CORAL 100% deductive system

Raghu Ramakrishnan , Divesh Srivastava , S. Sudarshan , Praveen Seshadri  
**The VLDB Journal — The International Journal on Very Large Data Bases** April 1994  
 Volume 3 Issue 2

CORAL is a deductive system that supports a rich declarative language, and an interface to C++, which allows for a combination of declarative and imperative programming. A CORAL declarative program can be organized as a collection of interacting modules. CORAL supports a wide range of evaluation strategies, and automatically chooses an efficient strategy for each module in the program. Users can guide query optimization by selecting from a wide range of control choices. The CORAL system provides ...

## 2 Space optimization in deductive databases

100%

Divesh Srivastava , S. Sudarshan , Raghu Ramakrishnan , Jeffrey F. Naughton  
**ACM Transactions on Database Systems (TODS)** December 1995  
 Volume 20 Issue 4

In the bottom-up evaluation of logic programs and recursively defined views on databases, all generated facts are usually assumed to be stored until the end of the evaluation. Discarding facts during the evaluation, however, can considerably improve the efficiency of the evaluation: the space needed to evaluate the program, the I/O costs, the costs of maintaining and accessing indices, and the cost of eliminating duplicates may all be reduced. Given an evaluation method that is sound, compl ...

### 3 Combinatorial pattern discovery for scientific data: some preliminary results 100%



Jason Tsong-Li Wang , Gung-Wei Chirn , Thomas G. Marr , Bruce Shapiro , Dennis Shasha , Kaizhong Zhang

**ACM SIGMOD Record , Proceedings of the 1994 ACM SIGMOD international conference on Management of data** May 1994

Volume 23 Issue 2

Suppose you are given a set of natural entities (e.g., proteins, organisms, weather patterns, etc.) that possess some important common externally observable properties. You also have a structural description of the entities (e.g., sequence, topological, or geometrical data) and a distance metric. Combinatorial pattern discovery is the activity of finding patterns in the structural data that might explain these common properties based on the metric. This paper presents an example o ...

### 4 Industry track papers: Mining heterogeneous gene expression data with time lagged recurrent neural networks 99%



Yulan Liang , Arpad Kelemen

**Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining** July 2002

Heterogeneous types of gene expressions may provide a better insight into the biological role of gene interaction with the environment, disease development and drug effect at the molecular level. In this paper for both exploring and prediction purposes a Time Lagged Recurrent Neural Network with trajectory learning is proposed for identifying and classifying the gene functional patterns from the heterogeneous nonlinear time series microarray experiments. The proposed procedures identify gene fun ...

### 5 A data-analysis pipeline for large-scale gene expression analysis 99%



S. Hennig , R. Herwig , M. Clark , P. Aanstad , A. Musa , J. O'Brien , C. Bull , U. Radelof , G. Panopoulou , A. J. Poustka , H. Lehrach

**Proceedings of the fourth annual international conference on Computational molecular biology** April 2000

In this article we describe a method for characterization of large cDNA clone libraries based on oligonucleotide fingerprints (OFPs). The main advantage of this technique lies in that, without sequencing, each clone is tagged in an almost unique way, which has a couple of interesting applications, e.g. clustering of clones that belong to the same gene or gene family followed by sequencing of representative clones for each cluster. Moreover, small clusters are likely to represent rarely expres ...

### 6 Application of intelligent agent technology for managerial data analysis and mining 99%



Ranjit Bose , Vijayan Sugumaran

**ACM SIGMIS Database** January 1999  
Volume 30 Issue 1

Data analysis and mining technologies help bring business intelligence into organizational decision support systems (DSS). While a myriad of data analysis and mining technologies are commercially available today, organizations are seeing a growing gap between powerful storage (data warehouse) systems and the business users' ability to analyze and act effectively on the information they contain. We contend that to narrow this gap effectively, a data analysis and mining environment is needed that ...

**7 Industrial/government track: Mining hepatitis data with temporal abstraction** 99%



Tu Bao Ho , Trong Dung Nguyen , Saori Kawasaki , Si Quang Le , Dung Duc Nguyen , Hideto Yokoi , Katsuhiko Takabayashi

**Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining** August 2003

The hepatitis temporal database collected at Chiba university hospital between 1982-2001 was recently given to challenge the KDD research. The database is large where each patient corresponds to 983 tests represented as sequences of irregular timestamp points with different lengths. This paper presents a temporal abstraction approach to mining knowledge from this hepatitis database. Exploiting hepatitis background knowledge and data analysis, we introduce new notions and methods for abstracting ...

**8 A multi-expert system for the automatic detection of protein domains from sequence information** 99%



Niranjana Nagarajan , Golan Yona

**Proceedings of the seventh annual international conference on Computational molecular biology** April 2003

We describe a novel method for detecting the domain structure of a protein from sequence information alone. The method is based on analyzing multiple sequence alignments that are derived from a database search. Multiple measures are defined to quantify the domain information content of each position along the sequence, and are combined into a single predictor using a neural network. The output is further smoothed and post-processed using a probabilistic model to predict the most likely transitio ...

**9 Analysis of the context dependency of CODASYL find-statements with application to a database program conversion** 98%



G. Barbara Demko , Sukhamay Kundu

**Proceedings of the 1985 ACM SIGMOD international conference on Management of data** May 1985

**10 Burst tries: a fast, efficient data structure for string keys** 97%



**ACM Transactions on Information Systems (TOIS)** April 2002

Volume 20 Issue 2

Many applications depend on efficient management of large sets of distinct strings in memory. For example, during index construction for text databases a record is held for each distinct word in the text, containing the word itself and information such as counters. We propose a new data structure, the burst trie, that has significant advantages over existing options for such applications: it uses about the same memory as a binary search tree; it is as fast as a trie; and, while not as fast as a ...

**11 XTRACT: a system for extracting document type descriptors from XML documents** 96%



Minos Garofalakis , Aristides Gionis , Rajeev Rastogi , S. Seshadri , Kyuseok Shim

**ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data** May 2000

Volume 29 Issue 2

XML is rapidly emerging as the new standard for data representation and exchange

on the Web. An XML document can be accompanied by a *Document Type Descriptor* (DTD) which plays the role of a schema for an XML data collection. DTDs contain valuable information on the structure of documents and thus have a crucial role in the efficient storage of XML data, as well as the effective formulation and optimization of XML queries. In this paper, we propose XTRACT, a novel system for inferring a ...

## 12 LogicBase: a deductive database system prototype

96%



Jiawei Han , Ling Liu , Zhaohui Xie

**Proceedings of the third international conference on Information and knowledge management** November 1994

A deductive database system prototype, LogicBase, has been developed, with an emphasis on efficient compilation and query evaluation of application-oriented recursions in deductive databases. The system identifies different classes of recursions and compiles recursions into chain or psuedo-chain forms when appropriate. Queries posed to the compiled recursions are analyzed systematically with efficient evaluation plans generated and executed, mainly based on a chained-based quer ...

## 13 Algorithms on Stings, Trees, and Sequences: Computer Science and

91%



Computational Biology

Dan Gusfield

**ACM SIGACT News** December 1997

Volume 28 Issue 4

## 14 Program Transformation Systems

90%



H. Partsch , R. Steinbrüggen

**ACM Computing Surveys (CSUR)** September 1983

Volume 15 Issue 3

## 15 Notung: dating gene duplications using gene family trees

89%



Kevin Chen , Dannie Durand , Martin Farach-Colton

**Proceedings of the fourth annual international conference on Computational molecular biology** April 2000

Large scale gene duplication is a major force driving the evolution of genetic functional innovation. Whole genome duplications are widely believed to have played an important role in the evolution of the maize, yeast and vertebrate genomes. The use of evolutionary trees to analyze the history of gene duplication and estimate duplication times provides a powerful tool for studying this process. Many studies in the molecular evolution literature have used this approach on small data sets, usin ...

## 16 Federated database systems for managing distributed, heterogeneous, and autonomous databases

89%



Amit P. Sheth , James A. Larson


**ACM Computing Surveys (CSUR)** September 1990

Volume 22 Issue 3


A federated database system (FDBS) is a collection of cooperating database systems that are autonomous and possibly heterogeneous. In this paper, we define a reference architecture for distributed database management systems from system and schema viewpoints and show how various FDBS architectures can be developed. We then define a methodology for developing one of the popular architectures of an

FDBS. Finally, we discuss critical issues related to developing and operating an FDBS.


**17** Managing conflicts between rules (extended abstract) 89%

 H. V. Jagadish , Alberto O. Mendelzon , Inderpal Singh Mumick  
**Proceedings of the fifteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems** June 1996


**18** Cost and availability tradeoffs in replicated data concurrency control 89%

 Akhil Kumar , Arie Segev  
**ACM Transactions on Database Systems (TODS)** March 1993  
 Volume 18 Issue 1

**19** Mining features for sequence classification 88%

 Neal Lesh , Mohammed J. Zaki , Mitsunori Ogihara  
**Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining** August 1999

**20** Disk cache—miss ratio analysis and design considerations 87%

 Alan J. Smith  
**ACM Transactions on Computer Systems (TOCS)** August 1985  
 Volume 3 Issue 3

The current trend of computer system technology is toward CPUs with rapidly increasing processing power and toward disk drives of rapidly increasing density, but with disk performance increasing very slowly if at all. The implication of these trends is that at some point the processing power of computer systems will be limited by the throughput of the input/output (I/O) system. A solution to this problem, which is described and evaluated in this paper, is disk cache

---

Results 1 - 20 of 200

short listing

  
 Prev  
 Page

1 2 3 4 5 6 7 8 9 10

  
 Next  
 Page

---

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2004 ACM, Inc.



[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent & Trademark Office



Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: **[((database or data <near> mining) and (sequence or gene) <near> analysis and (subset) <paragraph> (reduce or eliminate) <near> (duplicate or redundant))]**

Found **43** of **126,861** searched.

## Search within Results



[> Advanced Search](#)

[> Search Help/Tips](#)

Sort by: **Title** **Publication** **Publication Date** **Score** Binder

Results 1 - 20 of 43    short listing

Prev  
Page

1

2

3

Next  
Page

### 1 A relational approach to monitoring complex systems 87%



Richard Snodgrass

**ACM Transactions on Computer Systems (TOCS)** May 1988

Volume 6 Issue 2

Monitoring is an essential part of many program development tools, and plays a central role in debugging, optimization, status reporting, and reconfiguration. Traditional monitoring techniques are inadequate when monitoring complex systems such as multiprocessors or distributed systems. A new approach is described in which a historical database forms the conceptual basis for the information processed by the monitor. This approach permits advances in specifying the low-level data collection, ...

### 2 Knowledge discovery in data warehouses 87%



Themistoklis Palpanas

**ACM SIGMOD Record** September 2000

Volume 29 Issue 3

As the size of data warehouses increase to several hundreds of gigabytes or terabytes, the need for methods and tools that will automate the process of knowledge extraction, or guide the user to subsets of the dataset that are of particular interest, is becoming prominent. In this survey paper we explore the problem of identifying and extracting interesting knowledge from large collections of data residing in data warehouses, by using data mining techniques. Such techniques have the ability to i ...

### 3 Identifying the most significant pairwise correlations of residues in 85%



different positions of helices: the subset selection problem using least squares optimization

Xianghong Zhou , Gareth Chelvanayagam , Michael Hallett

**Proceedings of the 2001 ACM symposium on Applied computing** March 2001

- 4** Web crawling and measurement: Efficient URL caching for world wide web crawling 82%



Andrei Z. Broder , Marc Najork , Janet L. Wiener

**Proceedings of the twelfth international conference on World Wide Web** May 2003

Crawling the web is deceptively simple: the basic algorithm is (a) Fetch a page (b) Parse it to extract all linked URLs (c) For all the URLs not seen before, repeat (a)-(c). However, the size of the web (estimated at over 4 billion pages) and its rate of change (estimated at 7% per week) move this plan from a trivial programming exercise to a serious algorithmic and system design challenge. Indeed, these two factors alone imply that for a reasonably fresh and complete crawl of the web, step (a) ...

- 5** Response Time Analysis of Multiprocessor Computers for Database Support 81%



Roger K. Shultz , Roy J. Zingg

**ACM Transactions on Database Systems (TODS)** March 1984

Volume 9 Issue 1

Comparison of three multiprocessor computer architectures for database support is made possible through evaluation of response time expressions. These expressions are derived by parameterizing algorithms performed by each machine to execute a relational algebra query. Parameters represent properties of the database and components of the machines. Studies of particular parameter values exhibit response times for conventional machine technology, for low selectivity, high duplicate occurrence, ...

- 6** Beyond islands (extended abstract): runs in clone-probe matrices 77%



David B. Wilson , David S. Greenberg , Cynthia A. Phillips

**Proceedings of the first annual international conference on Computational molecular biology** January 1997

- 7** Articles: Data analysis and mining in the life sciences 72%



Nam Huyn

**ACM SIGMOD Record** September 2001

Volume 30 Issue 3

Biotech companies routinely generate vast amounts of biological measurement data that must be analyzed rapidly and mined for diagnostic, prognostic, or drug evaluation purposes. While these data analysis tasks are critical to their success, they have not benefited from recent advances that emerged from database and KDD research. In this paper, we focus on two such tasks: on-line analysis of clinical study data, and mining broad datasets for biomarkers. We examine the new requirements that are no ...

- 8** A taxonomy of parallel sorting 70%



Dina Bitton , David J. DeWitt , David K. Hsaio , Jaishankar Menon

**ACM Computing Surveys (CSUR)** September 1984

Volume 16 Issue 3

- 9** Scalable parallel data mining for association rules 69%



Eui-Hong Han , George Karypis , Vipin Kumar

**ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international**

**conference on Management of data June 1997**

Volume 26 Issue 2

One of the important problems in data mining is discovering association rules from databases of transactions where each transaction consists of a set of items. The most time consuming operation in this discovery process is the computation of the frequency of the occurrences of interesting subset of items (called candidates) in the database of transactions. To prune the exponentially large space of candidates, most existing algorithms, consider only those candidates that have a user defined ...

**10 A perspective on inductive databases**

64%



Luc De Raedt

**ACM SIGKDD Explorations Newsletter** December 2002

Volume 4 Issue 2

Inductive databases tightly integrate databases with data mining. The key ideas are that data and patterns (or models) are handled in the same way and that an inductive query language allows the user to query and manipulate the patterns (or models) of interest. This paper proposes a simple and abstract model for inductive databases. We describe the basic formalism, a simple but fairly powerful inductive query language, some basics of reasoning for query optimization, and discuss some memory organ ...

**11 Compilers 2: Compiler supported high-level abstractions for sparse disk-resident datasets**

64%



Renato Ferreira , Gagan Agrawal , Joel Saltz

**Proceedings of the 16th international conference on Supercomputing** June 2002

Processing and analyzing large volumes of data plays an increasingly important role in many domains of scientific research. The complexity and irregularity of datasets in many domains make the task of developing such processing applications tedious and error-prone. We propose use of high-level abstractions for hiding the irregularities in these datasets and enabling rapid development of correct data processing applications. We present two execution strategies and a set of compiler analysis techni ...

**12 Performance analysis in the software lifecycle: The Sisyphus database retrieval software performance antipattern**

61%



Robert F. Dugan , Ephraim P. Glinert , Ali Shokoufandeh

**Proceedings of the third international workshop on Software and performance** July 2002

In this paper we propose the Sisyphus database retrieval software performance antipattern. The antipattern occurs in application designs that process large, frequently accessed lists stored in a relational database, but display only a small subset to the user. Software Performance Engineering (SPE) techniques are used to analyze the antipattern. Four solutions are evaluated: rownum and index, upper/lower bound, sequence numbering, and caching. We discuss the real world challenges of correcting t ...

**13 Efficient clustering of high-dimensional data sets with application to reference matching**

57%



Andrew McCallum , Kamal Nigam , Lyle H. Ungar

**Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining** August 2000**14 Packet classification on multiple fields**

55%

Pankaj Gupta , Nick McKeown




-  **ACM SIGCOMM Computer Communication Review , Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication** August 1999  
Volume 29 Issue 4

Routers classify packets to determine which flow they belong to, and to decide what service they should receive. Classification may, in general, be based on an arbitrary number of fields in the packet header. Performing classification quickly on an arbitrary number of fields is known to be difficult, and has poor worst-case performance. In this paper, we consider a number of classifiers taken from real networks. We find that the classifiers contain considerable structure and redundancy that can ...

## 15 Declarative updates of relational databases


47%

-  Weidong Chen  
**ACM Transactions on Database Systems (TODS)** March 1995  
Volume 20 Issue 1

This article presents a declarative language, called update calculus, of relational database updates. A formula in update calculus involves conditions for the current database, as well as assertions about a new database. Logical connectives and quantifiers become constructors of complex updates, offering flexible specifications of database transformations. Update calculus can express all nondeterministic database transformations that are polynomial time. For set-a ...

## 16 Index scans using a finite LRU buffer: a validated I/O model

46%

-  Lothar F. Mackert , Guy M. Lohman  
**ACM Transactions on Database Systems (TODS)** September 1989  
Volume 14 Issue 3

Indexes are commonly employed to retrieve a portion of a file or to retrieve its records in a particular order. An accurate performance model of indexes is essential to the design, analysis, and tuning of file management and database systems, and particularly to database query optimization. Many previous studies have addressed the problem of estimating the number of disk page fetches when randomly accessing k records out of N given records stored on

## 17 A transformational approach to compiling Sisal for distributed memory architectures

35%

-  Michael O'Boyle , G. A. Hedayat  
**Proceedings of the 6th international conference on Supercomputing** August 1992

This paper is concerned with the efficient execution of array computation on Distributed Memory Architectures by applying compiler-directed program and data transformations. By translating a subset of a single-assignment language, Sisal, into a linear algebraic framework it is possible to transform a program so as to reduce load imbalance and non-local memory access. A new test is presented which allows the construction of transformations to reduce load imbalance. By a new expression of dat ...

## 18 The program decision logic approach to predicated execution

33%

-  David I. August , John W. Sias , Jean-Michel Puiatti , Scott A. Mahlke , Daniel A. Connors , Kevin M. Crozier , Wen-mei W. Hwu  
**ACM SIGARCH Computer Architecture News , Proceedings of the 26th annual international symposium on Computer architecture** May 1999  
Volume 27 Issue 2

Modern compilers must expose sufficient amounts of Instruction-Level Parallelism (ILP) to achieve the promised performance increases of superscalar and VLIW

processors. One of the major impediments to achieving this goal has been inefficient programmatic control flow. Historically, the compiler has translated the programmer's original control structure directly into assembly code with conditional branch instructions. Eliminating inefficiencies in handling branch instructions and exploiting ILP h ...

**19 RuleViz: a model for visualizing knowledge discovery process**

31%



Jianchao Han , Nick Cercone

**Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining** August 2000

**20 An intermediate design language and its analysis**

31%



Daniel Jackson

**ACM SIGSOFT Software Engineering Notes , Proceedings of the 6th ACM SIGSOFT international symposium on Foundations of software engineering** November 1998  
Volume 23 Issue 6

A simple relational language is presented that has two desirable properties. First, it is sufficiently expressive to encode, fairly naturally, a variety of software design problems. Second, it is amenable to fully automatic analysis. This paper explains the language and its semantics, and describes a new analysis scheme (based on a stochastic boolean solver) that dramatically outperforms existing schemes.

---

**Results 1 - 20 of 43      short listing**

  
Prev  
Page

**1   2   3**

  
Next  
Page

---

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2004 ACM, Inc.

[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent &amp; Trademark Office

Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: **[(annotation database and (sequence or biopolymer or dna or macromolecule or protein or nucleotides) <paragraph> (sequence or order or arrangement or classify or organize) <paragraph> (id or identifier or identification))]**

Found 3 of 126,861 searched.

091

## Search within Results

[> Advanced Search](#)[> Search Help/Tips](#)

---

**Sort by:** Title Publication Publication Date Score Binder

---

**Results 1 - 3 of 3** short listing

---

**1** A platform for the description, distribution and analysis of genetic 77%

polymorphism data

Greg D. Tyrelle , Garry C. King

**Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003 - Volume 19** January 2003

In this paper we suggest the requirements for an open platform designed for the description, distribution and analysis of genetic polymorphism data. This platform is discussed in terms of our implementation of a phenotypic prediction pipeline with general application to the understanding of genetic variation. The current state of polymorphism data storage and distribution has several recognised deficiencies. These include the lack of a shared data model and low overlap between databases. To move ...

**2** A new approach to protein structure and function analysis using semi- 77%

structured databases

William M. Shui , Raymond K. Wong , Stephen C. Graham , Lawrence K. Lee , W. Bret Church

**Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003 - Volume 19** January 2003

The development of high-throughput genome sequencing and protein structure determination techniques have provided researchers with a wealth of biological data. Integrated analysis of such data is difficult due to the disparate nature of the repositories used to store this biological data and of the software used for its analysis. This paper presents a framework based upon the use of semi-structured database management systems that would provide an integrated interface for the collection, storage ...

### 3 Genome scale prediction of protein functional class from sequence using 77% data mining



Ross D. King , Andreas Karwath , Amanda Clare , Luc Dephaspe

**Proceedings of the sixth ACM SIGKDD international conference on Knowledge  
discovery and data mining** August 2000

---

#### Results 1 - 3 of 3      short listing

---

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2004 ACM, Inc.



[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent & Trademark Office



Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: **[((database or data mining) and (sequence or biopolymer or dna or macromolecule or protein or nucleotides) <paragraph> (sequence or order or arrangement or classify or organize) <paragraph> (id or identifier or identification) and (reduce or eliminate) <paragraph> (duplicate or redundant))]**

Found **248** of **126,861** searched.

**Warning: Maximum result set of 200 exceeded. Consider refining.**

## Search within Results



[> Advanced Search](#)

[> Search Help/Tips](#)

**Sort by:** Title Publication Publication Date Score

**Results 1 - 20 of 200**

**short listing**



**1 2 3 4 5 6 7 8 9 10**



- 1** **Automatic high-quality reengineering of database programs by abstraction, transformation and reimplementation** 100%

Yossi Cohen , Yishai A. Feldman

**ACM Transactions on Software Engineering and Methodology (TOSEM)** July 2003  
Volume 12 Issue 3

Old-generation database models, such as the indexed-sequential, hierarchical, or network models, provide record-level access to their data, with all application logic residing in the hosting program. In contrast, relational databases can perform complex operations, such as filter, aggregation, and join, on multiple records without an external specification of the record-access logic. Programs written for relational databases attempt to move as much of the application logic as possible into the d ...
- 2** **Path sharing and predicate evaluation for high-performance XML filtering** 100%

Yanlei Diao , Mehmet Altinel , Michael J. Franklin , Hao Zhang , Peter Fischer

**ACM Transactions on Database Systems (TODS)** December 2003  
Volume 28 Issue 4

XML filtering systems aim to provide fast, on-the-fly matching of XML-encoded data to large numbers of query specifications containing constraints on both structure and content. It is now well accepted that approaches using event-based parsing and Finite State Machines (FSMs) can provide the basis for highly scalable structure-oriented XML filtering systems. The XFilter system [Altinel and Franklin 2000] was the first published FSM-based XML filtering approach. XFilter used a separate FSM per

pa ...

### 3 Efficient dynamic mining of constrained frequent sets 100%



Laks V. S. Lakshmanan , Carson Kai-Sang Leung , Raymond T. Ng  
**ACM Transactions on Database Systems (TODS)** December 2003

Volume 28 Issue 4

Data mining is supposed to be an iterative and exploratory process. In this context, we are working on a project with the overall objective of developing a practical computing environment for the human-centered exploratory mining of frequent sets. One critical component of such an environment is the support for the dynamic mining of constrained frequent sets of items. Constraints enable users to impose a certain focus on the mining process; dynamic means that, in the middle of the computation, u ...

### 4 Query optimization in a memory-resident domain relational calculus database system 100%



Kyu-Young Whang , Ravi Krishnamurthy  
**ACM Transactions on Database Systems (TODS)** March 1990

Volume 15 Issue 1

We present techniques for optimizing queries in memory-resident database systems. Optimization techniques in memory-resident database systems differ significantly from those in conventional disk-resident database systems. In this paper we address the following aspects of query optimization in such systems and present specific solutions for them: (1) a new approach to developing a CPU-intensive cost model; (2) new optimization strategies for main-memory query processing; (3) new insight into ...

### 5 Description logics for semantic query optimization in object-oriented database systems 100%



Domenico Beneventano , Sonia Bergamaschi , Claudio Sartori  
**ACM Transactions on Database Systems (TODS)** March 2003

Volume 28 Issue 1

Semantic query optimization uses semantic knowledge (i.e., integrity constraints) to transform a query into an equivalent one that may be answered more efficiently. This article proposes a general method for semantic query optimization in the framework of Object-Oriented Database Systems. The method is effective for a large class of queries, including conjunctive recursive queries expressed with regular path expressions and is based on three ingredients. The first is a Description Logic, ODL

### 6 A logic for object-oriented logic programming 100%



M. Kifer , J. Wu  
**Proceedings of the eighth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems** March 1989

We present a logic for reasoning about complex objects, which is a revised and significantly extended version of Maier's O-logic [Mai86]. The logic naturally supports complex objects, object identity, deduction, is tolerant to inconsistent data, and has many other interesting features. It elegantly combines the object-oriented and value-oriented paradigms and, in particular, contains all of the predicate calculus as a special case. Our treatment of sets is also noteworthy: it is more genera ...

### 7 A temporally oriented data model 100%



Gad Ariav

**ACM Transactions on Database Systems (TODS)** December 1986

Volume 11 Issue 4

The research into time and data models has so far focused on the identification of extensions to the classical relational model that would provide it with "adequate" semantic capacity to deal with time. The temporally oriented data model (TODM) presented in this paper is a result of a different approach, namely, it directly operationalizes the pervasive three-dimensional metaphor for time. One of the main results is thus the development of the notion of the data cube: a three-di ...

**8 A perspective on inductive databases**

100%



Luc De Raedt

**ACM SIGKDD Explorations Newsletter** December 2002

Volume 4 Issue 2

Inductive databases tightly integrate databases with data mining. The key ideas are that data and patterns (or models) are handled in the same way and that an inductive query language allows the user to query and manipulate the patterns (or models) of interest. This paper proposes a simple and abstract model for inductive databases. We describe the basic formalism, a simple but fairly powerful inductive query language, some basics of reasoning for query optimization, and discuss some memory organ ...

**9 Parallelizing OODBMS traversals: a performance evaluation**

100%



David J. De Witt , Jeffrey F. Naughton , John C. Shafer , Shivakumar Venkataraman

**The VLDB Journal — The International Journal on Very Large Data Bases** January 1996

Volume 5 Issue 1

In this paper we describe the design and implementation of *ParSets*, a means of exploiting parallelism in the SHORE OODBMS. We used *ParSets* to parallelize the graph traversal portion of the OO7 OODBMS benchmark, and present speedup and scaleup results from parallel SHORE running these traversals on a cluster of commodity workstations connected by a standard ethernet. For some OO7 traversals, SHORE achieved excellent speedup and scaleup; for other OO7 traversals, only marginal speedup and s ...

**10 Concurrency and recovery for index trees**

100%



David Lomet , Betty Salzberg

**The VLDB Journal — The International Journal on Very Large Data Bases** August 1997

Volume 6 Issue 3

Although many suggestions have been made for concurrency in B<sup>+</sup>-trees, few of these have considered recovery as well. We describe an approach which provides high concurrency while preserving well-formed trees across system crashes. Our approach works for a class of index trees that is a generalization of the B<sup>+</sup><sub>link</sub>-tree. This class includes some multi-attribute indexes and temporal indexes. Structural changes in an index tree are decomposed into a sequence of atomic actions, each one ...

**11 Special issue on persistent object systems: Fibonacci: a programming language for object databases**

100%



Antonio Albano , Giorgio Ghelli , Renzo Orsini

**The VLDB Journal — The International Journal on Very Large Data Bases** July 1995

Volume 4 Issue 3

Fibonacci is an object-oriented database programming language characterized by static and strong typing, and by new mechanisms for modeling databases in terms of objects with roles, classes, and associations. A brief introduction to the language is provided to present those features, which are particularly suited to modeling complex databases. Examples of the use of Fibonacci are given with reference to the prototype implementation of the language.

## 12 Special issue on prototypes of deductive database systems: The CORAL 100% deductive system



Raghu Ramakrishnan , Divesh Srivastava , S. Sudarshan , Praveen Seshadri

**The VLDB Journal — The International Journal on Very Large Data Bases** April 1994

Volume 3 Issue 2

CORAL is a deductive system that supports a rich declarative language, and an interface to C++, which allows for a combination of declarative and imperative programming. A CORAL declarative program can be organized as a collection of interacting modules. CORAL supports a wide range of evaluation strategies, and automatically chooses an efficient strategy for each module in the program. Users can guide query optimization by selecting from a wide range of control choices. The CORAL system provides ...

## 13 Column: Generating consistent test data: restricting the search space 100% by a generator formula



Andrea Neufeld , Guido Moerkotte , Peter C. Lockemann

**The VLDB Journal — The International Journal on Very Large Data Bases** April 1993

Volume 2 Issue 2

To address the problem of generating test data for a set of general consistency constraints, we propose a new two-step approach: First the interdependencies between consistency constraints are explored and a generator formula is derived on their basis. During its creation, the user may exert control. In essence, the generator formula contains information to restrict the search for consistent test databases. In the second step, the test database is generated. Here, two different approaches are pr ...

## 14 Maintaining availability in partitioned replicated databases 100%



A. El Abbadi , S. Toueg

**ACM Transactions on Database Systems (TODS)** June 1989

Volume 14 Issue 2

In a replicated database, a data item may have copies residing on several sites. A replica control protocol is necessary to ensure that data items with several copies behave as if they consist of a single copy, as far as users can tell. We describe a new replica control protocol that allows the accessing of data in spite of site failures and network partitioning. This protocol provides the database designer with a large degree of flexibility in deciding the degree of data availability, as w ...

## 15 SilkRoute: A framework for publishing relational data in XML 100%



Mary Fernández , Yana Kadiyska , Dan Suciu , Atsuyuki Morishima , Wang-Chiew Tan

**ACM Transactions on Database Systems (TODS)** December 2002

Volume 27 Issue 4

XML is the "lingua franca" for data exchange between interenterprise applications. In this work, we describe SilkRoute, a framework for publishing relational data in XML. In SilkRoute, relational data is published in three steps: the relational tables are



presented to the database administrator in a canonical XML view; the database administrator defines in the XQuery query language a public, virtual XML view over the canonical XML view; and an application formulates an XQuery query over the publ ...

## 16 Modeling the storage architectures of commercial database systems 100%



D. S. Batory

**ACM Transactions on Database Systems (TODS)** December 1985

Volume 10 Issue 4

Modeling the storage structures of a DBMS is a prerequisite to understanding and optimizing database performance. Previously, such modeling was very difficult because the fundamental role of conceptual-to-internal mappings in DBMS implementations went unrecognized. In this paper we present a model of physical databases, called the transformation model, that makes conceptual-to-internal mappings explicit. By exposing such mappings, we show that it is possible to model the storage ...

## 17 Data conversion and restructuring: An Access Path Specification 100%



Language for restructuring network databases

Donald Swartwout

**Proceedings of the 1977 ACM SIGMOD international conference on Management of data** August 1977

The Access Path Specification Language (APSL) is a high-level essentially nonprocedural language for specifying restructuring transformations of network databases. It does so in terms of application-oriented concepts such as access strategies and selection criteria. APSL's approach to restructuring emphasizes description of the source *structures* from which target data is to be retrieved, rather than the *operations* needed to convert source constructs to target constructs. The latter ...

## 18 Modeling concepts for VLSI CAD objects 100%



D. S. Batory , Won Kim

**ACM Transactions on Database Systems (TODS)** September 1985

Volume 10 Issue 3

VLSI CAD applications deal with design objects that have an interface description and an implementation description. Versions of design objects have a common interface but differ in their implementations. A molecular object is a modeling construct which enables a database entity to be represented by two sets of heterogeneous records, one set describes the object's interface and the other describes its implementation. Thus a reasonable starting point for modeling design objects is to begin w ...

## 19 Distributed query processing 100%



C. T. Yu , C. C. Chang

**ACM Computing Surveys (CSUR)** December 1984

Volume 16 Issue 4

## 20 Implementation concepts for an extensible data model and data 100%



language

D. S. Batory , T. Y. Leung , T. E. Wise

**ACM Transactions on Database Systems (TODS)** September 1988

Volume 13 Issue 3

Future database systems must feature extensible data models and data languages in

order to accommodate the novel data types and special-purpose operations that are required by nontraditional database applications. In this paper, we outline a functional data model and data language that are targeted for the semantic interface of GENESIS, an extensible DBMS. The model and language are generalizations of FQL [11] and DAPLEX [40], and have an implementation that fits ideally with the modularity ...

---

**Results 1 - 20 of 200****short listing**Prev  
Page**1 2 3 4 5 6 7 8 9 10**Next  
Page

---

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2004 ACM, Inc.